

# OPTIMAL EMBEDDINGS AND EIGENVALUES IN SUPPORT THEORY

ERIK G. BOMAN\*, STEPHEN GUATTERY†, AND BRUCE HENDRICKSON\*

**Abstract.** Support theory is a methodology for bounding eigenvalues and generalized eigenvalues of matrices and matrix pencils; such bounds have been stated both in algebraic terms, and in combinatorial terms based on embeddings of the underlying graphs of the matrices. In this paper, we present a theorem that demonstrates the connection between these various bounding techniques, and also suggests a possible approach to generating approximate inverses for preconditioning. The theorem shows, given matrices  $A = UD_AU^*$  and  $B = VD_BV^*$  (where  $D_A$  and  $D_B$  are invertible Hermitian matrices, and  $U$  and  $V$  are not necessarily square), that it is possible to define a matrix  $W$  such that  $W^*D_B^{-1}WD_A$  has the same nonzero eigenvalues counting multiplicity as  $B^+A$ . In the special case that  $U$  is the orthogonal projector onto the range space of  $B$  and  $D_A = I$  (and hence that  $A = UU^* = U^2 = U$ ), then  $W^*D_B^{-1}W = B^+$ . This suggests that finding an approximation to  $W$  might lead to an approximate inverse that can be used in preconditioning. We also describe how this theorem generalizes the idea of graph embeddings in an algebraic sense.

**1. Introduction.** Support theory is a methodology for bounding eigenvalues and generalized eigenvalues (and more generally support numbers) of matrices and matrix pencils; it has applications such as the analysis of the performance of iterative solvers for symmetric positive definite systems. Support theory bounds have been stated both in algebraic terms, and in terms of combinatorial techniques based on the underlying graphs of the matrices involved. In this paper, we present a theorem that demonstrates the connection between these various bounding techniques, and also suggests a possible approach to generating approximate inverses for preconditioning.

Given a preconditioned system  $B^{-1}A$ , support theory is concerned with generating bounds on  $\lambda_{\max}(B^{-1}A)$  and  $\lambda_{\max}(A^{-1}B)$ , which are used to bound the condition number of the system. More generally, if  $A$  and  $B$  are singular, bounds are generated on the support numbers  $\sigma(A, B)$  and  $\sigma(B, A)$ . Bounds are generated in terms of factorizations  $A = UU^T$  and  $B = VV^T$  of the matrices where  $U$  and  $V$  are not restricted to be square. Given the factorizations, a matrix  $W$  is constructed subject to the condition  $U = VW$ . This is the key formulation of support theory [3].

We present a theorem that shows, given matrices  $A = UD_AU^*$  and  $B = VD_BV^*$  (where  $D_A$  and  $D_B$  are invertible Hermitian matrices and the columns of  $U$  are in the range of  $B$ ), that there exists a matrix  $W_{opt}$  such that  $W_{opt}^*D_B^{-1}W_{opt}D_A$  has the same nonzero eigenvalues counting multiplicity as  $B^+A$ . This result was previously unknown even in the special case where  $D_A = D_B = I$ . In the special case that  $U$  is the orthogonal projector of the range space of  $B$  and  $D_A = I$  (and hence that  $A = UU^* = U^2 = U$ ), then  $W_{opt}^*D_B^{-1}W_{opt} = B^+$ , where  $B^+$  is the pseudoinverse of  $B$ . This suggests that finding an approximation to  $W_{opt}$  might lead to an approximate inverse that can be used in preconditioning.

We describe how this theorem generalizes the idea of embedding in an algebraic sense, and show how it can be used to generalize and simplify the proofs of previous

---

\*Discrete Algorithms and Math Dept, Sandia National Laboratories, Albuquerque, NM 87185-1111. {egboman,bahendr}@sandia.gov. Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin company, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

†Department of Computer Science, Bucknell University, Lewisburg, PA 17837. guattery@bucknell.edu. Parts of this work were carried out while the author was on a sabbatical appointment with the Computer Science Research Institute, Sandia National Labs, NM.

results. In particular, this allows the ideas to be applied to a broader range of matrices. We also place the theorem in the context of two branches of support theory research that focus on matrix eigenvalue bounds and matrix pencil bounds respectively.

The paper starts with a section on notation (Section 2) and some background on the development of some basic ideas in support theory (Section 3). The main theorem is presented in Section 4, along with a version applicable to non-Hermitian matrices. We use this proof to give a new and more general proof of a result by Boman and Hendrickson [3]. In Section 6 we give new and more general proofs of results by Guattery [6] linking a generalized notion of embedding with the pseudoinverses of Hermitian matrices. The results in this paper subsume the results in that tech report. Finally, in Section 7 we discuss the implications of this work for preconditioning, particularly in terms of approximate inverses.

**2. Notation.** We use capital letters to represent matrices. Individual matrix entries are denoted by the corresponding lower case letter with subscripts showing the row and column of the entry, e.g.,  $a_{ij}$  is the entry of  $A$  in row  $i$  and column  $j$ .  $I$  represents an identity matrix. When it is useful to indicate the size of an identity matrix, a single subscript indicates the number of rows and columns:  $I_k$  is the  $k \times k$  identity matrix. A matrix of all zeros is denoted by  $0$ .

For a matrix  $A$  with real entries,  $|A|$  denotes the matrix whose entries are the absolute values of the corresponding entries of  $A$ : the entry in row  $i$  and column  $j$  of  $|A|$  is  $|a_{ij}|$ .

The notations  $A^T$  represents the transpose of  $A$ ;  $A^*$  represents  $A$ 's conjugate transpose; and  $A^+$  represents the pseudoinverse of  $A$  (i.e., the Moore-Penrose generalized inverse of  $A$ ).

We denote the range space of the columns of a matrix  $A$  by  $R(A)$ . The orthogonal projector onto a vector space  $\mathcal{S}$  is denoted as  $P_{\mathcal{S}}$ . Thus  $P_{R(A)}$  is the orthogonal projector onto the range space of the columns of  $A$ .

Vectors are denoted by lower case letters with an arrow above, e.g.  $\vec{v}$ . A column vector of all zeros is denoted as  $\vec{0}$ ; if a specific size is specified, it is given as a subscript:  $\vec{0}_k$  is a vector of  $k$  0's.

**3. Background.** A key application of support theory is the analysis of preconditioned symmetric and Hermitian systems. When solving linear systems  $Ax = b$  using an iterative method, it is frequently useful to have a good preconditioner to accelerate convergence. This has often involved constructing a preconditioner  $B \approx A$ . In this sense,  $B$  is a good preconditioner if both (i) the eigenvalues of  $B^{-1}A$  are clustered around one, and (ii) the matrix  $B$  is easy to solve for (invert). (Note that if  $A$  is singular, one may wish to let  $B$  be singular with the same null space. In this case,  $B^+A$  is the preconditioned matrix of interest.)

For Hermitian, positive definite systems, the eigenvalues are real and positive, and a lower bound on the rate of convergence is the spectral condition number,  $\kappa(C) = \lambda_{\max}(C)/\lambda_{\min}(C)$ . In our case, either  $C = B^{-1}A$  or  $C = B^{-1/2}AB^{-1/2}$  (where  $B^{-1/2}$  is Hermitian), and the condition number can be expressed using support numbers. The *support number* for a matrix pencil  $(A, B)$ , where  $A$  and  $B$  are Hermitian, is defined as

$$\sigma(A, B) = \min \{t \in \mathbb{R} \mid x^*(\tau B - A)x \geq 0 \text{ for all } x \in \mathbb{C}^n \text{ and for all } \tau \geq t\}.$$

It has been shown that when  $B$  is symmetric positive definite,  $\sigma(A, B) = \lambda_{\max}(A, B)$ ,

the largest generalized eigenvalue [3]. Furthermore,

$$\kappa(B^{-1/2}AB^{-1/2}) = \sigma(A, B)\sigma(B, A).$$

Support numbers exist even when  $A$  or  $B$  is singular, but may not be finite. Support theory is useful for analyzing preconditioners because support numbers give bounds on the spectral condition number. See [3] for further information on support theory.

For symmetric positive semidefinite systems, bounds on support numbers are generated in terms of factorizations  $A = UU^T$  and  $B = VV^T$  of the matrices. Based on the factorizations, a matrix  $W$  is constructed subject to the condition  $U = VW$ . This key algebraic formulation of support theory is due to Boman and Hendrickson [3], and is at the heart of the symmetric product support theorem:

**THEOREM 3.1** (Theorem 4.5 from [3]). *Suppose  $U \in \mathbb{R}^{n \times k}$  is in the range of  $V \in \mathbb{R}^{n \times p}$ . Then*

$$\sigma(UU^T, VV^T) = \min_W \|W\|_2^2 \text{ subject to } VW = U.$$

It is further noted in [3] that for  $\hat{W} = V+U$ ,  $\sigma(UU^T, VV^T) = \|\hat{W}\|_2^2$ .

They also proved the following:

**THEOREM 3.2** (Theorem 4.7 from [3]). *Suppose  $U \in \mathbb{R}^{n \times k}$  is in the range of  $V \in \mathbb{R}^{n \times p}$ , and  $D \in \mathbb{R}^{k \times k}$  is symmetric. Then for all  $W$  such that  $VW = U$ ,*

$$\sigma(UDU^T, VV^T) \leq \lambda_{\max}(WDW^T) \leq \lambda_{\max}(D)\|W\|_2^2.$$

The use of the term *embedding* in this paper is intended to suggest the relationship to the use of combinatorial embeddings in early work in the area, which typically involved Laplacian (and related) matrices. Laplacians have a well-known graphical interpretation in which a Laplacian  $L$  can be factored into  $L = UU^T$ , where  $U$  is a *vertex-edge incidence matrix*. In such a matrix, each row represents a vertex of the underlying graph of the Laplacian. Each column represents an edge, and has one  $+1$  and one  $-1$  in the rows corresponding to the endpoints of the edge. Note that this (arbitrarily) directs the edges, though this direction is lost in the product that forms the Laplacian. (Entries can be scaled appropriately for weighted graphs.)

Graph embedding techniques were used to analyze the quality of support graph preconditioners, which were formulated in terms of the combinatorial structure of the Laplacian matrices. Examples include Vaidya's preconditioners based on spanning trees [13] (see [2] for a more readily available description of Vaidya's preconditioner), and Gremban and Miller's support tree preconditioners [5]. Bounds on the condition number of a preconditioned system were calculated in terms of properties of the embeddings of the underlying graphs of the preconditioner into the Laplacian and vice versa.

The embeddings used in this analysis were typically path embeddings, in which each edge in  $U$  (e.g., the vertex-edge incidence matrix of the original Laplacian) was represented as a directed path constructed from the edges in  $V$  (e.g., the vertex-edge incidence matrix of the preconditioner). While these embeddings were seldom not typically expressed in matrix form, it is easy to construct the matrix  $W$  corresponding to the embedding: Each path in the embedding shows up as a column in  $W$ . The rows in  $W$  correspond to the edges in  $V$ ; if an edge in  $V$  is in the path, a  $+1$  or  $-1$

occurs in that row of  $W$ , depending on whether the direction of the edge in  $V$  is in the same or opposite direction of the path respectively. (Again, entries in  $W$  can be scaled to deal with weighted graphs.) In this representation,  $U = VW$  expresses the mapping defined by the embedding.

Properties of the paths were used to compute bounds. This was typically done in combinatorial rather than algebraic terms. One widely used method, suggested in Vaidya's work (see the section *An analogy with resistive networks* on p. 6 of [13]) and developed by Gremban and Miller (see the discussion starting with the last two paragraphs on p. 65 and continuing to the start of Section 4.5 on the next page of [5]), involved summing the congestions (defined for each edge in the graph embedded into as the number of paths in the embedding incident to that edge) along each path. The maximum sum taken over all paths provided an upper bound on  $\lambda_{\max}(B^{-1}A)$ .

This method was also applied in a line of research that considered bounding the smallest nontrivial eigenvalue of a Laplacian or related matrix (the earliest work in this direction was aimed at bounding the second largest eigenvalues of time-reversible Markov chains in order to bound the mixing time for random walks [10, 12]). This typically involved embedding the complete graph into a Laplacian or a generalization of the Laplacian that allowed weighted edges. Later work involved a further generalization that included Dirichlet boundary conditions (these conditions resulted in entries on the diagonal that exceeded the sum of the weights of the incident edges). For such matrices, a star was embedded instead of a clique.

Analysis using path congestions and related techniques initially did not express the embedding in matrix form. However, some work along these lines developed an algebraic representation of embeddings. Kahale [11] looked at computing lower bounds on the smallest nontrivial eigenvalue of a Laplacian using a method that assigned a length to each path, then looked for the edge that had the greatest sum of the lengths of the incident paths. He computed embedding properties in terms of  $|W|$ , the absolute value of the matrix  $W$  representing the embedding, and expressed the best bound that could be computed for any embedding in terms of  $\lambda_{\max}(|W||W|^T)$ . Guattery, Leighton, and Miller [7] formulated the path resistance method, an extension of the sum-of-congestions method applied to lower bounds of smallest nontrivial Laplacian eigenvalues. They showed that the path-resistance method was a dual of Kahale's edge-length method in that both had the same best possible bound given some embedding  $W$ , and that the value  $\lambda_{\max}(|W||W|^T)$  was a term in the expression for this best bound.

Guattery and Miller [8] made the observation that including directions in embeddings typically improved the best possible lower bounds on smallest nontrivial eigenvalues of generalized Laplacians that could be derived from the embeddings. They allowed multiple paths in embedding each edge. They also kept the signs corresponding to direction in the embedding matrix  $W$ , which corresponds to working with  $W$  rather than  $|W|$ . At this point the notion of embedding has been generalized to the point that any matrix  $W$  such that  $U = VW$  can be viewed as a *generalized embedding* of  $U$  into  $V$ , though  $U$  and  $V$  are still vertex-edge incidence matrices. Guattery and Miller also proved that there existed a particular embedding  $W_{cf}$  (referred to as the *current flow embedding* such that the smallest nontrivial eigenvalue could be computed exactly in terms of  $\lambda_{\max}(W_{cf}W_{cf}^T)$  for that embedding. They also showed that  $W_{cf}$  was a factor of the inverse for this embedding applied to the case of a generalized Laplacian with a Dirichlet boundary condition.

Guattery [6] extended these ideas to all Hermitian matrices. In particular, he

considered what he called the *generalized current flow embedding*  $W_{cf}$  of the orthogonal projector onto the range of the columns of  $B = VDV^*$  (where  $D$  is a diagonal matrix with nonzeros on the diagonal) into  $V$ . He showed that the pseudoinverse of  $B$  could be expressed as  $W_{cf}D^{-1}W_{cf}^*$ . The formulation is still in terms of a slightly generalized vertex-edge incidence matrix although, as shown in Section 6, the ideas can be generalized to any factorization  $VDV^*$ .

The results in this paper tie together ideas from the algebraic and combinatorial (path embedding) views of support theory, and express them in a common notation.

**4. The Main Theorem.** Consider the matrix pencil  $(A, B)$ , where  $A$  and  $B$  are  $n \times n$  complex, Hermitian matrices. Assume that there exist matrices  $U$  and  $V$  such that  $A = UD_AU^*$  and  $B = VD_BV^*$ , where  $U$  is  $n \times m$ ,  $V$  is  $n \times k$ , and  $D_B$  and  $D_A$  are invertible Hermitian matrices with dimensions  $k \times k$  and  $m \times m$  respectively. The main theorem can be stated as follows:

**THEOREM 4.1 (Main Theorem).** *Given the matrix pencil  $(A, B)$  and factorizations described above, if the columns of  $U$  are in the range of  $B$ , then there exists a matrix  $W_{opt}$  such that  $VW_{opt} = U$ , and such that every nonzero eigenvalue of  $W_{opt}^*D_B^{-1}W_{opt}D_A$  is also an eigenvalue of  $B^+A$ , and vice versa, counting multiplicity.*

Before proving the main theorem, it is helpful to lay out some supporting lemmas. Define the  $(n+k) \times (n+k)$  block matrix  $M$  as follows:

$$M = \begin{bmatrix} D_B^{-1} & V^* \\ V & 0 \end{bmatrix}. \quad (4.1)$$

The proof of the main theorem is based on solutions to the following block system:

$$\begin{bmatrix} D_B^{-1} & V^* \\ V & 0 \end{bmatrix} \begin{bmatrix} W_{opt} \\ Z \end{bmatrix} = \begin{bmatrix} 0 \\ U \end{bmatrix}. \quad (4.2)$$

The  $k \times m$  matrix  $W_{opt}$  is of particular interest. The following lemma gives necessary and sufficient conditions for the existence of a solution, and hence of  $W_{opt}$ :

**LEMMA 4.2.** *The system above has a solution if and only if the columns of  $U$  are in the range of  $B$ .*

*Proof.* We can apply Gaussian elimination to  $M$  as follows:

$$\begin{bmatrix} I & 0 \\ -VD_B & I \end{bmatrix} \begin{bmatrix} D_B^{-1} & V^* \\ V & 0 \end{bmatrix} \begin{bmatrix} W_{opt} \\ Z \end{bmatrix} = \begin{bmatrix} I & 0 \\ -VD_B & I \end{bmatrix} \begin{bmatrix} 0 \\ U \end{bmatrix}$$

to get the following reduced system:

$$\begin{bmatrix} D_B^{-1} & V^* \\ 0 & -B \end{bmatrix} \begin{bmatrix} W_{opt} \\ Z \end{bmatrix} = \begin{bmatrix} 0 \\ U \end{bmatrix}. \quad (4.3)$$

Proving the result for the reduced matrix is sufficient because it has the same solutions as the original system.

Assuming that a solution to the reduced system exists, there must be a  $Z$  such that  $-BZ = U$ . That is, a solution exists only if the columns of  $U$  are in the range of  $B$ .

Assuming that the columns of  $U$  are in the range of  $B$ , then a  $Z$  that satisfies  $-BZ = U$  exists. For the block system to have a solution, we must also have a solution to the equation  $D_B^{-1}W_{opt} + V^*Z = 0$ . The solution exists if there is a  $W_{opt}$

such that  $W_{opt} = -D_B V^* Z$ . Since the matrices  $D_B$ ,  $V^*$ , and  $Z$  exist by assumption, a solution does exist, and the lemma holds.  $\square$

It is possible that there is more than one solution. In some arguments below we require a solution that is orthogonal to the null space of  $M$ ; where this is the case, it is noted explicitly.

Theorem 1.3.20 from reference [9] is useful in our arguments. We restate (in slightly revised form) the pertinent result below:

**THEOREM 4.3.** *Let  $Y$  be an  $r \times s$  matrix and  $Z$  be an  $s \times r$  matrix, with  $r \leq s$ . Then  $ZY$  has the same eigenvalues as  $YZ$ , counting multiplicity, together with an additional  $s - r$  eigenvalues equal to 0.*

We can now prove the main theorem:

*Proof.* [Proof of Main Theorem] Given the factorizations assumed in the theorem statement, we can construct the block system in equation 4.2. As in Lemma 4.2, we can apply Gaussian elimination to get the reduced system from equation (4.3). Assume that the columns of  $U$  are in the range of  $B$ . By Lemma 4.2, a solution to the reduced system exists. That immediately yields the equation

$$-BZ = U.$$

Multiplying through on both sides by  $-B^+$  gives

$$P_{R(B^+)}Z = -B^+U.$$

It is a fact that  $P_{R(B^+)} = P_{R(B^*)}$  (see p. 10 of reference [4]). Thus we can rewrite the equation as follows:

$$P_{R(B^*)}Z = -B^+U. \quad (4.4)$$

The assumption that the columns of  $U$  are in the range of  $B$  also implies that  $U = P_{R(B)}U$ . Taking the conjugate transpose gives  $U^* = U^*P_{R(B)}^* = U^*P_{R(B)}$ , the last equality following because orthogonal projectors are Hermitian by definition. Because  $B$  is Hermitian,  $P_{R(B)} = P_{R(B^*)}$ . Thus we have the following equation:

$$U^* = U^*P_{R(B^*)}. \quad (4.5)$$

From the original block system we have that

$$D_B^{-1}W_{opt} = -V^*Z, \quad (4.6)$$

and also that, after taking the conjugate transpose,

$$W_{opt}^*V^* = U^*. \quad (4.7)$$

Applying equations (4.6), (4.7), (4.5) and (4.4) in turn, we have

$$\begin{aligned} W_{opt}^*D_B^{-1}W_{opt}D_A &= -W_{opt}^*V^*ZD_A \\ &= -U^*ZD_A \\ &= -U^*P_{R(B^*)}ZD_A \\ &= U^*B^+UD_A. \end{aligned} \quad (4.8)$$

We can apply Theorem 4.3 to show that each nonzero eigenvalue of  $U^*B^+UD_A$  is an eigenvalue of  $B^+A$  and vice-versa:

If  $m > n$ , we apply Theorem 4.3 to  $B^+A = B^+UD_AU^*$ . Noting that  $B^+UD_A$  is  $n \times m$  and  $U^*$  is  $m \times n$ , we immediately have that  $B^+A$  has the same nonzero eigenvalues (counting multiplicity) as  $U^*B^+UD_A$ , plus  $m - n$  additional zero eigenvalues.

If  $m \leq n$ , we apply Theorem 4.3 to  $U^*B^+UD_A$ . We immediately have that  $B^+UD_AU^* = B^+A$  has the same nonzero eigenvalues as  $U^*B^+UD_A$  (counting multiplicity), plus  $n - m$  additional zero eigenvalues if  $n$  is strictly greater than  $m$ .

This completes the proof.

□

As a consequence of Theorem 4.1, we have the additional result:

**THEOREM 4.4.** *Suppose  $U \in \mathbb{R}^{n \times k}$  is in the range of  $V \in \mathbb{R}^{n \times p}$ . Then there exists a  $W$  such that  $VW = U$  and  $W^TW$  has the same nonzero eigenvalues as  $(VV^T)^+UU^T$ .*

*Proof.* The result follows immediately from Theorem 4.1, with  $D_B = I$  and  $D_A = I$ . □

Note that in this case,  $D_B = I$  in the system from equation 4.2, so in the solution orthogonal to the null space of  $M$  (Equation 4.1),  $W$  is the solution to the equation  $VW = U$  with the minimum two-norm. This is consistent with Theorem 3.1 (Theorem 4.5 from [3]).

**5. The Non-Hermitian Case.** A related theorem can be proved for invertible, complex, non-Hermitian matrices of the form  $A = EF^*$ , where  $A$  is  $n \times n$ , and  $E$  and  $F$  are  $n \times m$ , with  $m \geq n$ . Consider the following two block systems, one for  $A$ :

$$\begin{bmatrix} I_m & F^* \\ E & 0 \end{bmatrix} \begin{bmatrix} W \\ Z \end{bmatrix} = \begin{bmatrix} 0 \\ I_n \end{bmatrix}, \quad (5.1)$$

and one for  $A^* = FE^*$ :

$$\begin{bmatrix} I_m & E^* \\ F & 0 \end{bmatrix} \begin{bmatrix} X \\ Y \end{bmatrix} = \begin{bmatrix} 0 \\ I_n \end{bmatrix}. \quad (5.2)$$

With respect to these systems and the assumption that  $A$  is invertible, we have the following theorem:

**THEOREM 5.1.**  $X^*W = A^{-1}$ .

*Proof.* The system in equation (5.1) yields the equations

$$W = -F^*Z$$

and

$$EW = I_n, \quad (5.3)$$

which combined yield the equation

$$-EF^*Z = I_n,$$

which implies that

$$-Z = A^{-1}.$$

Likewise the system in equation (5.2) yields the equations

$$X = -E^*Y, \quad (5.4)$$

$$FX = I_n,$$

$$-FE^*Y = I_n,$$

and

$$-Y = (A^*)^{-1}. \quad (5.5)$$

Hence

$$\begin{aligned} X^*W &= -Y^*EW && \text{(by (5.4))} \\ &= -Y^* && \text{(by (5.3))} \\ &= A^{-1} && \text{(by (5.5))} \end{aligned}$$

□

**6. Generalized Embeddings and Pseudoinverses.** The main theorem can also be applied to producing pseudoinverses. Guattery, in a technical report [6], showed that by applying a generalized version of embedding to the orthogonal projector onto the range space of any symmetric matrix, it is possible to generate factors of that matrix's pseudoinverse. He also showed how to extend this result to all Hermitian matrices by splitting them into a real and an imaginary part, and working with a system twice as big. Theorem 4.1 allows us to prove the results from this technical report in a simpler way. The new proof covers the Hermitian case directly.

Let  $B$  be a Hermitian matrix, with  $B = VDV^*$ ,  $D$  Hermitian and invertible. Let  $\begin{bmatrix} W \\ Z \end{bmatrix}$  be the solution to the system

$$\begin{bmatrix} D^{-1} & V^* \\ V & 0 \end{bmatrix} \begin{bmatrix} W \\ Z \end{bmatrix} = \begin{bmatrix} 0 \\ P_{R(B)} \end{bmatrix}, \quad (6.1)$$

where  $P_{R(B)}$  is the orthogonal projector onto the range space of  $B$ . The matrix  $W$  here corresponds to what Guattery termed a *generalized embedding*.

The following two theorems are direct consequences of Theorem 4.1 and its proof:  
THEOREM 6.1.  $W^*D^{-1}W = B^+$ .

*Proof.* Viewing the system above in terms of the system in Theorem 4.1, we can make the substitutions:  $U = P_{R(B)}$ ,  $D_B = D$ , and  $D_A = I_n$ . Substituting these values in equation (4.8) from the proof of Theorem 4.1 gives

$$W^*D^{-1}W = P_{R(B)}^*B^+P_{R(B)}.$$

Note that  $P_{R(B^+)} = P_{R(B^*)}$  (see reference [4]). Recall that  $P_{R(B)} = P_{R(B^*)}$  since  $B$  is Hermitian, and that  $P_{R(B)}^* = P_{R(B)}$  because orthogonal projectors are Hermitian by definition. These equations also imply that  $P_{R(B^+)} = P_{R(B)}$ . Hence we have the following:

$$P_{R(B)}^*B^+P_{R(B)} = P_{R(B^+)}B^+P_{R(B)} = B^+P_{R(B)} = B^+.$$

The last equality follows from the fact that  $B^+$  is Hermitian, so  $R(B^+)$  is a reducing subspace. Proposition 0.2.3 from reference [4] thus gives us that

$$P_{R(B^+)}B^+ = B^+P_{R(B^+)} = B^+.$$



Since  $P_{R(B^+)} = P_{R(B)}$  (as noted above), this proves the theorem.

□

In the proof above, note that the factor embedded,  $P_{R(B)}$ , is also equal to the symmetric product of the factor and its conjugate transpose. This is done to insure  $W$  has the proper dimensions to serve as a factor of the pseudoinverse.

One possibility hidden in the argument is that the rank of  $B$  may be less than the rank of  $V$ . In such cases the matrix  $D$  projects part of the range of  $V^*$  into the null space of  $V$ . When this does not happen, however, we can prove an interesting property. Let matrices  $V$ ,  $D$ , and  $B$  be defined as for Theorem 6.1. Let  $\begin{bmatrix} W \\ Z \end{bmatrix}$  be a solution to equation 6.1 that is orthogonal to the null space of the block matrix. We have the following:

**THEOREM 6.2.** *If  $\text{rank}(B) = \text{rank}(V)$ ,  $-Z = B^+$ .*

*Proof.* Assume that the order of  $D$  is  $k$ , and that  $V$  is an  $n \times k$  matrix.

We can apply Gaussian elimination to the system in equation 6.1 as follows:

$$\begin{bmatrix} I & 0 \\ -VD & I \end{bmatrix} \begin{bmatrix} D^{-1} & V^* \\ V & 0 \end{bmatrix} \begin{bmatrix} W \\ Z \end{bmatrix} = \begin{bmatrix} I & 0 \\ -VD & I \end{bmatrix} \begin{bmatrix} 0 \\ P_{R(B)} \end{bmatrix}$$

to get the following block upper triangular system:

$$\begin{bmatrix} D^{-1} & V^* \\ 0 & -B \end{bmatrix} \begin{bmatrix} W \\ Z \end{bmatrix} = \begin{bmatrix} 0 \\ P_{R(B)} \end{bmatrix},$$

The rank of the block matrix (referred to as  $M$  below) is greater than or equal to the sum of the ranks of  $B$  and  $D^{-1}$  (see, e.g., reference [9]). Since  $D^{-1}$  has full rank, this also implies that the dimension of the null space of  $M$  is less than or equal to the size of the null space of  $B$ .

Note that we have  $-BZ = P_{R(B)}$ .

If  $B$  is nonsingular, then the result follows immediately:  $P_{R(B)} = I$ , and  $-BZ = I$ . By the uniqueness of the inverse,  $Z = -B^{-1}$ .

If  $B$  has less than full rank, multiplying through on both sides by  $B^+$  gives the equation

$$-P_{R(B)}Z = B^+P_{R(B)} = B^+,$$

where the last equality follows by the argument given in Theorem 6.1. The theorem therefore holds if the columns of  $Z$  are in the range of  $B$ .

Since  $B = VDV^*$ , the null space of  $V^*$  is contained in the null space of  $B$ . The condition in the theorem statement that the ranks of  $B$  and  $V$  are the same thus implies that the null spaces of  $B$  and  $V^*$  are the same.

We can construct a basis for the null space of the block matrix  $M$  as follows: Assume the size of the null space of  $B$  is  $j \geq 1$ . Construct an orthogonal basis for the null space of  $B$  consisting of vectors  $\vec{v}_1 \dots \vec{v}_j$ . Consider the  $j$  vectors of the form  $\vec{w}_i = \begin{bmatrix} \vec{0}_k \\ \vec{v}_i \end{bmatrix}$ , where  $\vec{0}_k$  is the column vector consisting of  $k$  zeros. These vectors are clearly orthogonal. Because  $B$  and  $V^*$  have the same null spaces, each such vector is in the null space of  $M$ . And because (as noted above) the size of the null space of  $M$  is less than or equal to size of the null space of  $B$ , the  $\vec{w}_i$ 's span the null space of  $M$ , and hence form a basis for it.

By assumption, the solution  $\begin{bmatrix} W \\ Z \end{bmatrix}$  is orthogonal to the null space of  $M$ . By the structure of the vectors in the null space of  $M$ , this means that  $Z$  is orthogonal to the null space of  $B$ . Hence  $Z$  is in the range of  $B$ , and the theorem holds.

□

**7. Approximate Inverse Preconditioning.** In preconditioning linear systems  $Ax = b$ , one often constructs a preconditioner  $B \approx A$ . In the iterative method, one has to solve for (invert)  $B$ . An alternative is to algebraically construct a matrix  $M$  such that  $MA \approx I$ . Since  $M$  approximates the inverse of  $A$ , one only needs to multiply by  $M$  in the iterative method, which has certain advantages (e.g., in parallel computing).

Many strategies have been proposed for constructing approximate inverses, see e.g. [1]. One practical condition is that  $M$  must be sparse. Some try to minimize  $\|MA - I\|$  (with certain sparsity constraints) directly, while others construct  $M$  as the product of two triangular factors.

An interesting open question is whether our Theorem 6.1 can be used as a starting point for constructing approximate inverses. This theorem provides a novel factorization for the inverse of a symmetric matrix. Are there cases in which we could compute an inexpensive approximation  $\tilde{W}$  to  $W$ ? Then  $\tilde{W}^* D^{-1} \tilde{W}$  would be an approximation to the inverse.

The key issues to be settled are whether there exist classes of matrices for which a good approximation to  $W$  exists. Additionally, there is the question of which factorization of  $A$  (if any) can be used to find good approximate factors, and how the approximation is constructed based on such a factorization.

It is also be interesting to consider a similar strategy applied to nonsymmetric matrices using Theorem 5.1. Can good approximations to the factors  $X$  and  $W$  of our nonsymmetric matrix can be generated by approximately solving Equations 5.1 or 5.2?

We remark that approximate inverses of this type are in factored form but the factors may be rectangular (non-square), which would make such an approach distinctively different from other factored approximate inverses.

## REFERENCES

- [1] M. BENZI AND M. TUMA, *A comparative study of sparse approximate inverse preconditioners*, Applied Numerical Mathematics, 30 (1999), pp. 305–340.
- [2] M. BERN, J. GILBERT, B. HENDRICKSON, N. NGUYEN, AND S. TOLEDO, *Support-graph preconditioners*, in Proceedings of the 5th Copper Mountain Conference on Iterative Methods, Copper Mountain, Colorado, 1998, pp. 1–7.
- [3] E. G. BOMAN AND B. HENDRICKSON, *Support theory for preconditioning*, SIAM J. on Matrix Anal. and Appl., 25 (2004), pp. 694–717. (Published electronically on 17 Dec 2003.).
- [4] S. L. CAMPBELL AND C. D. MEYER, JR., *Generalized Inverses of Linear Transformations*, Dover, New York, 1991. Reprint of a work originally published by Pitman Publishing Limited (London) in 1979.
- [5] K. D. GREMBAN, *Combinatorial Preconditioners for Sparse, Symmetric, Diagonally Dominant Linear Systems*, PhD thesis, Carnegie Mellon University, Pittsburgh, October 1996. CMU CS Tech Report CMU-CS-96-123.
- [6] S. GUATTERY, *Graph embeddings, symmetric real matrices, and generalized inverses*, Tech. Report 98-34, ICASE, August 1998.
- [7] S. GUATTERY, F. T. LEIGHTON, AND G. L. MILLER, *The path resistance method for bounding the smallest nontrivial eigenvalue of a laplacian*, Combinatorics, Probability and Computing, 8 (1999), pp. 441–460.

- [8] S. GUATTERY AND G. L. MILLER, *Graph embeddings and Laplacian eigenvalues*, SIAM J. Mat. Anal. App., 21 (2000), pp. 703–723.
- [9] R. A. HORN AND C. A. JOHNSON, *Matrix Analysis*, Cambridge University Press, Cambridge, 1985.
- [10] M. JERRUM AND A. SINCLAIR, *Approximating the permanent*, SIAM Journal on Computing, 18 (1989), pp. 1149–1178.
- [11] N. KAHALE, *A semidefinite bound for mixing rates of Markov chains*, in Proc. Fifth Integer Programming and Combinatorial Optimization Conference, 1996, pp. 190–203. Lecture Notes in Computer Science 1084.
- [12] A. SINCLAIR AND M. JERRUM, *Approximate counting, uniform generation and rapidly mixing Markov chains*, Information and Computation, 82 (1989), pp. 93–133.
- [13] P. M. VAIDYA, *Solving linear equations with symmetric diagonally dominant matrices by constructing good preconditioners*. Unpublished manuscript presented at the IMA Workshop on Graph Theory and Sparse Matrix Computation, October 1991.